

ВЕЧНЫЕ ВОПРОСЫ ПО ПОВОДУ МЕХАНИЗМОВ КАТЕГОРИЗАЦИИ: «СКОЛЬКО?», «КАКИЕ?», «ЗАЧЕМ?»

Н.П. Радчикова,

кандидат психологических наук, доцент
Белорусский государственный педагогический
университет имени Максима Танка

Проблема категоризации появилась в 70-х годах благодаря исследованиям Э. Рош и до сих пор вызывает ожесточенные споры. От единой концепции классификации, основанной на существенных признаках, научное сообщество обратилось к вероятностным моделям (теория прототипов, экземпляров и пр.), основанным на сходстве, которое определялось различными способами. Вскоре, однако, обнаружилось, что и эти модели имеют существенные недостатки, и их общность была поставлена под сомнение. Научное сообщество оказалось в положении, когда одна группа теорий объясняет один перечень явлений и данных, а вторая – другой [1,2]. Совершенно естественно, что возникла идея существования, по крайней мере, двух механизмов формирования и использования категорий.

Первую попытку соединить вместе два механизма представляет собой теория сравнения признаков (feature comparison theory) [3], предполагающая, что есть два типа признаков, по которым осуществляется сравнение родовой и видовой категории – существенные и характеристические. Существенные признаки определяют членство в категории, а характеристические – типичность каждого ее члена. Процесс отнесения объекта к какой-либо категории происходит в два этапа. Сначала сравниваются все признаки (и существенные, и характеристические), а на втором этапе – только существенные. Например, при определении, является ли воробей птицей, происходит совпадение многих признаков воробья и птицы, поэтому процесс категоризации завершается быстро. При определении того, является ли пингвин птицей, первый этап сравнения дает мало совпадающих характеристических признаков, поэтому следует перейти ко второму этапу – сравнению существенных признаков. В этом случае время реакции будет больше. Такое предположение является дополнением и исправлением классической теории категоризации. Однако оно объясняет только процесс верификации предложений типа «Воробей – это птица?» либо механизм опознавания объекта.

К основной критике этой теории относят то, что она не дает ответа на вопрос, каким образом признаки делятся на существенные и характеристические. Многие исследователи сомневаются в том, что существенные признаки вообще существуют [4,5]. Следующей проблемой

является проблема четких границ между категориями, обычно упоминающаяся противниками классической теории категоризации (более подробное рассмотрение этой проблемы можно найти в [1,2]) Еще один аргумент против теории сравнения признаков – результаты исследования Э. Лофтус [6] Она измеряла продуктивную частотность категорий двумя способами: задавая испытуемым написать членов категории при предъявлении названия категории (член-для-категории) либо задавая испытуемым написать название категории при предъявлении членов категории (категория-для-члена). В соответствии с теорией сравнения признаков время реакции для вопросов «Воробей – это птица?» и «Птица – это воробей?» должно быть одинаковым. Однако оказалось, что продуктивная частотность членов-для-категории является хорошим предсказателем времени реакции для ответов на вопрос, в котором член стоит перед названием категории («Воробей – это птица?»). В то же время продуктивная частотность категорий-для-членов является хорошим предсказателем времени реакции для ответов на вопрос, в котором название категории стоит перед членом категории («Птица – это воробей?»).

Предлагались и другие варианты моделей, сходных с теорией сравнения признаков. E. Smith и D. Medin [7] утверждали, что есть некое ядро понятия и его процедура идентификации. Ядро состоит из существенных признаков, а узнавание объекта часто осуществляется по характеристическим признакам.

Если обратиться к более поздним работам, то можно привести следующие примеры, представляющие собой гораздо более формализованные модели.

Модель «Правило-плюс-Исключение» (RULEX), предложенная R. Nosofsky [8,9] предполагает, что для категоризации человек использует сначала простые одномерные правила. Если такая стратегия не работает (например, человек осознает, что делает слишком много ошибок), то происходит переход к более сложной комбинации простых правил. Если и эта стратегия не работает, то человек запоминает несколько исключений в дополнение к правилам. Таким образом, данная модель объединяет предположения классической теории и теории экземпляров. В целом, модель дает довольно хорошие предсказания эмпирических данных, однако проблема заключается в том, что она не объясняет эффекты трудно классифицируемых стимулов. Тем не менее, можно предположить, что выбор стратегии (простые правила, сложные правила или сложные правила+исключения) может меняться в зависимости от сложности распознавания стимулов.

Модель ATRIUM [10] имеет два модуля, которые работают параллельно. Один модуль основан на правилах, второй – на сравнении с примерами (экземплярами). При каждом новом акте категоризации оба модуля вычисляют активацию категории, а затем некий блок принятия решения определяет

влияние каждого модуля на категориальное решение. Этот блок является сложным, и его работа зависит от характера предъявляемых стимулов. Для стимулов, которые правильно классифицируются на основании правила, блок отдает предпочтение модулю, основанному на правилах. Для стимулов, которые правильно классифицируются на основании сходства с хранящимся в памяти экземпляром, блок отдает предпочтение модулю, основанному на сравнении с экземплярами. И оба модуля, и блок принятия решения обучаются в процессе категоризации. Стратегия подбора правила эффективнее, так как если правило верное, то оно позволяет категоризировать несколько последующих стимулов сразу. Запоминание экземпляров требует больше времени, усилий и ресурсов. В модели ATRIUM обучение соревновательно. Выбор стратегии определения правила приводит к уменьшению способности запоминать отдельные экземпляры, и наоборот. Данная модель хорошо предсказывает многие полученные данные, но содержит очень много параметров, и не ясно, как подбор этих параметров может определять ее предсказательную способность. Кроме того, модель ATRIUM явно не определяет, когда должно происходить изменение в выборе стратегии категоризации при изменении трудности распознавания стимулов.

Г. Эшби [11-13] предлагает модель COVIS, в которой две системы (имплицитная и эксплицитная) работают параллельно в процессе категоризации, хотя изначально доминирует эксплицитная система, основанная на правилах. Имплицитная система никак не осознается. Она работает по принципу более медленного ассоциативного научения. Г. Эшби реализовал свою модель и как нейрофизиологическую, и как компьютерную. Обе реализации предполагают соревнование между системами. Существует несколько подтверждений такой идеи. Например, известно, что время обучения зависит от того, можно ли сформулировать явное правило [14], что данная модель предсказывает. Также оказалось, что пациенты с повреждениями различных частей мозга имеют трудности с различными типами категорий, что может свидетельствовать о существовании, по крайней мере, двух различных систем категоризации [12,13]. С помощью нейрофизиологических методов установлено, что в головном мозге активизируются различные области в зависимости от конкретной задачи категоризации [15,16]. Тем не менее, F.G. Ashby и S.W. Ell [12,13] утверждают, что можно осознанно помнить несколько конкретных пар член-категория, что приводит к добавлению в систему третьего механизма.

Когнитивную психологию часто критикуют за неумеренное умножение механизмов, блоков и частей любой системы, поэтому признание двух, а то и трех совершенно различных механизмов категоризации требует не только солидного эмпирического обоснования, но и ответа на вопрос об экологической

валидности. Можно ли обойтись одной системой или все же их несколько? Введение двух систем можно, конечно, начинать с того, что одна никак не объясняет все имеющиеся данные (и подавляющее большинство исследований исходят из этого), однако более обоснованным представляется подход, в котором теоретически обоснована необходимость существования более одной системы.

В. Слуцкий [17,18] предположил, что в природе существует два типа категорий (статистически плотные, не обладающие ни одним существенным признаком, и статистически неплотные, обладающие существенными признаками), и для их формирования у человека должны были развиваться две различные системы. Это предположение явно неприемлемо, так как категории являются порождением психики человека, и существовать без познающего субъекта не могут. Поэтому такой подход не объясняет, почему нам не хватает одной системы.

Логическую необходимость наличия двух принципиально разных механизмов можно увидеть в признании структуры когнитивного бессознательного, согласно которой только гетерогенная, состоящая из нескольких независимых когнитивных подсистем, система способна производить проверку и сопоставление входящей информации [19]. Так как мы не можем сравнивать наши представления с реальностью, а можем сравнивать только представления с представлениями, то по принципу независимой проверяемости возникает необходимость существования хотя бы двух систем формирования этих самых представлений (возможно, и более двух). Согласно предложению В.М. Аллахвердова, эти подсистемы работают параллельно, а качественные результаты их работы сравниваются, давая возможность отследить совпадение либо несовпадение результатов. Механизмы этих двух подсистем должны быть принципиально разными (например, аналоговым и дискретным). Такая гипотеза объясняет наличие двух независимых систем (механизмов), участвующих в процессе категоризации: системы оценивания сходства (объединения объектов по большой совокупности характеристических признаков, холистической, аналоговой, ассоционистской) и системы разбиения, деления (основанной на правилах, дискретной).

Принимая во внимание то, что оба механизма работают параллельно, то всегда идет выдвижение и проверка гипотез. Также всегда идет непрерывное обучение. Если существенные признаки можно выделить, то они выделяются одним механизмом (основанным на правилах), и далее категоризация идет быстро и беспроblemно, так как второй механизм не противоречит первому. Работа второго механизма (основанного на сходстве) может выражаться в эффекте типичности, когда существует некий набор характеристических признаков, которые согласуются с существенными.

От модели ATRIUM данное предположение отличается отсутствием соревнования систем (модулей). Блок, в котором информация сравнивается, является простым. Он дает обратную связь о совпадении либо несовпадении ответов, и его работа не зависит от предъявляемых стимулов. Кроме того, аналоговая система не запоминает отдельные экземпляры.

Рассмотрим, как предложенная В.М. Аллахвердовым модель структуры когнитивного бессознательного объясняет процессы категоризации, и какие предсказания можно сделать, исходя из такого предположения.

1) Возможно, что такая модель может объяснить диссоциацию между определением сходства и категоризацией, достаточно широко представленную литературе [20-22]. Неизвестно, какой именно механизм используется при категоризации, и какой механизм используется, когда люди выполняют задание определения сходства.

2) Подтверждение модели, состоящей из двух систем, можно увидеть в результатах проверки теорий формирования понятий – теории прерывности и теории непрерывности. Теория непрерывности, основанная на ассоционистской модели, предсказывала S-образную кривую научения, так как любой признак объекта рассматривался как стимул, который получает либо не получает подкрепление на каждом последующем шаге обучения. При предъявлении многочисленных объектов сильнее всего будут подкрепляться только существенные признаки. Теория прерывности утверждает, что человек выдвигает гипотезы о правиле категоризации, а затем проверяет их. Следовательно, кривая научения будет представлять собой ломаную, у которой число правильных ответов сначала колеблется в пределах 50%, а затем резким скачком достигает 100%. Большинство экспериментальных результатов подтвердили теорию прерывности [23], но следует отметить, однако, что в экспериментах всегда использовались так называемые неплотные категории, всегда существовало правило (существенные признаки), по которому можно было отнести объект в одну или другую категорию.

Плотные категории не изучались в процессе проверки теорий прерывности и непрерывности. Однако, можно предположить, что в данном случае должна наблюдаться плавная кривая научения, так как механизм поиска существенных признаков даст сбой.

3) Исследования показали, что при формировании понятий смена правила классификации в то время, когда испытуемый еще не догадался об этом правиле, не приводит к увеличению времени обучения понятиям при использовании неплотных категорий [24, 25]. Этот результат также интерпретировался как свидетельство того, что человек выдвигает и проверяет гипотезы в процессе формирования понятий. Теория непрерывности предсказывала более длительный этап обучения в этом случае, так как уже

образовались бы некоторые ассоциации, мешающие усвоению правильного понятия. Для плотных категорий экспериментов не проводилось. Так как при их формировании должен быть задействован «аналоговый» механизм, то смена правила должна приводить к увеличению времени формирования понятия.

Таким образом, теория когнитивного бессознательного, предложенная В.М. Аллахвердовым, позволяет провести эмпирическую проверку некоторых выводов. В этом издании представлены две работы. Статья Е.А. Киштымовой посвящена дальнейшему исследованию диссоциации между задачами определения сходства и категоризации. В статье А.С. Карбалевич проверяется гипотеза о различных формах кривых научения в случае предъявления плотных и неплотных категорий.

РЕПОЗИТОРИЙ БГПУ

ЛИТЕРАТУРА

1. Лобанов, А.П. Проблема категоризации в когнитивной психологии / А.П. Лобанов, Н.П. Радчикова // Белорусский психологический журнал. – 2005. – № 1. – С. 13-19.
2. Радчикова, Н.П. Проблемы категоризации в логике и психологии: от Аристотеля до Рош // Н.П. Радчикова, А.П. Лобанов // Психологический журнал. – № 1 (13), 2007. – С. 30-35.
3. Smith, E.E. Structure and process in semantic memory: A featural model for semantic decisions / E.E. Smith, E.J. Shoben, L.J. Rips // Psychological Review. – 1974. – Vol. 81. – P. 214-241.
4. Wittgenstein, L. Philosophical investigations / L.Wittgenstein. – New York: Macmillan, 1953.
5. Hampton, J.A. Psychological representation of concepts / J.A. Hampton // Cognitive models of memory / Ed. M.A. Conway. – Hove: Psychology Press, 1997. – P. 81-110.
6. Loftus, E.F. Category dominance, instance dominance, and categorization time / E.F. Loftus // Journal of Experimental Psychology. – Vol. 97(1). – 1973. – P. 70-74.
7. Smith, E.E. Categories and concepts / E.E. Smith, D.L. Medin – Harvard University Press, 1981. – 204 p.
8. Nosofsky, R.M. Rule-plus-exception model of classification learning / R.M. Nosofsky, T.J. Palmeri, S.C. McKinley // Psychological Review. – 1994 – No. 101. – P. 53–79.
9. Nosofsky, R.M. A rule-plus-exception model for classifying objects in continuous-dimension spaces / R.M. Nosofsky, T.J. Palmeri // Psychonomic Bulletin & Review. – 1985. – No. 5. – P. 345–369.
10. Erickson, M.A. Rules and exemplars in category learning / M.A. Erickson, J.K. Kruschke // Journal of Experimental Psychology: General. – 1998. – No. 127. – P. 107–140.
11. Ashby, F.G. A neuropsychological theory of multiple systems in category learning / F.G. Ashby, L.A. Alfonso-Reese, A.U. Turken, E.M. Waldron // Psychological Review. – 1998. – No. 105. – P. 442–481.
12. Ashby, F.G. The neurobiology of human category learning / F.G. Ashby, S.W. Ell // Trends in Cognitive Sciences. – 2001. – No. 5. – P. 204–210.
13. Ashby, F.G. Single versus multiple systems of learning and memory / F.G. Ashby, S.W. Ell // In H. Pashler & J. Wixted (Eds.) Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology (3rd ed.). – New York: Wiley, 2002. – P. 655–691.
14. Maddox, W.T. Comparing decision bound and exemplar models of categorization / W.T. Maddox, F.G. Ashby // Perception & Psychophysics. – 1993. – No. 53. – P. 49–70.

15. Patalano, A.L. PET evidence for multiple strategies of categorization / A.L. Patalano, E.E. Smith, J. Jonides, R.A. Koeppel // *Cognitive, Affective & Behavioral Neuroscience*. – 2002. – No. 1. – P. 360–370.
16. Reber, P.J. Contrasting cortical activity associated with category memory and recognition memory / P.J. Reber, C.E.L. Stark, L.R. Squire // *Learning & Memory*. – 1998. – No. 5. P. 420–428.
17. Kloos, H. What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories / H. Kloos, V.M. Sloutsky // *Journal of Experimental Psychology: General*. – 2008. – No. 137(1). – P. 52–72.
18. Sloutsky, V.M. From Perceptual Categories to Concepts: What Develops? / V.M. Sloutsky // *Cognitive Science*. – 2010. – No. 34. – P. 1244–1286.
19. Allakhverdov, V.M. Does Consciousness exist? In What Sense? / V.M. Allakhverdov, V.A. Gershkovich // *Integrative Psychological and Behavioral Science* – 2010. – No. 44. – P. 340–347.
20. Rips, L.J. Similarity, typicality and categorization / L.J. Rips // *Similarity and analogical reasoning* / Eds. S. Vosniadou, A. Ortony. – Cambridge: Cambridge University Press, 1989. – P. 21–59.
21. Thibaut, J.-P. Dissociations between categorization and similarity judgements as a result of learning feature distributions / J.-P. Thibaut, M. Dupont, P. Anselme // *Memory & Cognition*, 2002. – No 30(4). – P. 647–656.
22. Киштымова, Е.А. Роль сходства в процессе категоризации / Е.А. Киштымова, Н.П. Радчикова // *Вести БГПУ*. – № 2. – 2013. – С. 36–41.
23. Bower, G. Concept identification / G. Bower, T. Trabasso // In R.C. Atkinson (ed.) *Studies in Mathematical Psychology*. – Stanford University Press, 1964. – 418 p.
24. Kendler, H.H., D'Amato, M.F. A comparison of reversal shifts and nonreversal shifts in human concept formation behavior / H.H. Kendler, M.F. D'Amato // *Journal of Experimental Psychology*. – 1955. – Vol. 49(3). – P. 165–174.
25. Bower, G. Reversals prior to solution in concept identification / G. Bower, T. Trabasso // *Journal of Experimental Psychology*. – 1963. – Vol. 66, No. 4. – P. 409–418